A Theory of Cyclogenesis Forced by Diabatic Heating.  
Part I: A Quasi-geostrophic Approach

YUH-LANG LIN

Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, North Carolina

(Manuscript received 5 December 1988, in final form 8 May 1989)

ABSTRACT

A quasi-geostrophic theory of cyclogenesis forced by a low-level diabatic heating in a backsheared baroclinic flow is proposed. Existence of wind reversal in one direction of the basic flow is an essential criterion to obtain a forced baroclinic wave in the vicinity of the heating region. It was found that an analogy exists for a quasi-geostrophic flow over a mountain and a region of steady-state diabatic heating. The relationship is described by $h(x, y) = (-g/T_0 N^2)T(x, y)$, where $h(x, y)$ is the mountain shape and $T(x, y)$ is the temperature anomaly.

The response of a backsheared baroclinic flow over a region of two-dimensional diabatic heating (cooling) is a coupled low-high (high-low) pressure pair located in the vicinity and on the downstream side of the heating (cooling), respectively. Physically, the growth of the pressure perturbation can be explained by a group velocity argument. The disturbance remains locally in the vicinity of the forcing due to zero phase speed of the forced baroclinic wave. The upshear phase tilt indicates that the disturbance is a baroclinic wave generated by diabatic forcing.

The response of an east-west backsheared baroclinic flow over an isolated region of diabatic heating with circular contours is a growing cyclone located near the center of the heat source. A coupled high pressure forms downstream of the diabatic heating. The disturbance is confined in a shallow layer. The forced low resembles the geometry of the heat source in the early stage. One interesting finding is that an inverted ridge forms downstream of the low.

When applied to East Coast cyclogenesis, a cyclone develops near the center of the region of maximum diabatic heating, i.e., near the western boundary of the Gulf Stream. The cutoff low remains in the vicinity of the diabatic heat source. Two regions of weaker high pressure form to the southeast and northwest corners of the low. To the south of the low, there exists a strong anticyclonic circulation. A confluent zone forms to the northeast of the low, while a diffusive zone forms to the southwest of the low. The genesis region and the flow pattern of the cyclone predicted by the theory are consistent with observations. With an easterly wind at the surface, the inverted trough-ridge couplet is more pronounced than with a northeasterly wind. The low starts to decay as it moves out of the concentrated heating region. The cyclone is produced hydrostatically by the less dense air above the heating region with the modification of the baroclinic effects.

1. Introduction

Cyclogenesis along the east coast of the United States has received considerable attention since the recent completion of the Genesis of Atlantic Low Experiments (GALE). These cyclones often form off the Carolina coast, develop rapidly, and move northeastward, which may bring heavy snowfall and damage over the mid-Atlantic states. Different kinds of approaches, such as observational data analysis (e.g., Bosart 1981; Uccellini et al. 1984; Forbes et al. 1987), numerical simulations (e.g., Anthes et al. 1983; Atlas 1987; Danard and Ellenton 1980; Nuss and Anthes 1987; Orlanski and Katzfey 1987) and theoretical studies (e.g., Smith 1986, hereafter S86), have been used to investigate the problem of East Coast cyclogenesis. A number of formation and/or deepening mechanisms for East Coast cyclones have been proposed: (i) secondary cyclone formation along the coastal front (Bosart et al. 1972), (ii) cold-air damming on the east side of the Appalachians (Richstein 1980), (iii) boundary layer processes (e.g., coastal frontogenesis, cold-air damming, the development of a low-level jet streak) enhanced by the development of an upper-level jet streak (Uccellini and Kokin 1987), (iv) lee cyclogenesis associated with a basic baroclinic flow (S86), (v) slantwise convection (Emmanuel 1979, personal communication). Notice that the above mechanisms are not necessarily exclusive. However, the physical mechanisms that control East Coast cyclogenesis are still not fully understood.

Due to the complicated topographical factors in this region, such as the orographic forcing, the low-level differential heating, the shape of the coastline, the landsea frictional effects, and the latent heating, the dominant mechanism may differ for different types of East...
Coast cyclogenesis. For example, the cold-air damming by the Appalachians may play an important role in type B cyclogenesis (Miller 1946) because this type of cyclone is often preceded by a cold wedge along the east side of the mountains (Richwein 1980). Observational studies (e.g., Miller 1946) also indicate that the paths of surface low pressure centers associated with type B cyclone are discontinuous across the Appalachians. That is, a new cyclone forms to the east of the mountains as the parent cyclone impinges upon the mountains from the west or southwest. Thus, the lee cyclogenesis may also be important in this type of development (S86). In type A cyclogenesis, the cyclone forms as a wave on the cold front associated with a cold synoptic-scale anticyclone located to the north of the Carolina coast. The 1979 Presidents’ Day snowstorm (Bosart 1981; Uccellini et al. 1984) and the cyclogenesis during second GALE intensive observation period (IOP 2; GALE 1986) are good examples. Under such a synoptic setting, the horizontal basic wind normal to the Carolina coastline reverses at about 850 mb as the onshore lower-tropospheric flow underlay the westertlies in the middle and upper atmosphere. For a flow normal to the coastline, there exists a wind reversal level and a continuous low-level differential heating associated with the strong temperature contrast of the cold continent and warm Gulf Stream. There is little doubt that the differential heating plays an important role in this type of cyclogenesis. However, the exact mechanism in forming the cyclone still needs to be investigated.

One of the pioneer works on cyclogenesis is Sutcliffe’s development theory (Sutcliffe 1947). The theory stated that the advection of vorticity aloft and heating between the surface and the level of nondivergence are two major contributions to surface cyclogenesis. In deriving a diagnostic relation between vertical velocity and horizontal advections of vorticity and temperature, Sutcliffe came very close to discovering the quasi-geostrophic theory. In the development theory, surface wind speed had been assumed to be negligible compared with thermal wind. However, this type of wind profile is rarely observed in East Coast cyclogenesis events. Significant contributions to our understanding of cyclogenesis had been advanced by Charmey (1947) and Eady (1949) in developing quasi-geostrophic instability theory. Inspired by the work of Sutcliffe, Petterssen (1956) proposed that the development of a surface depression often occurs when a slowly moving surface trough is overtaken by an upper-level rapidly moving trough. However, the cyclonic development may cease as the vorticity configuration evolves at a later time. Using the quasi-geostrophic surface height tendency equation, Petterssen et al. (1962) found that the prediction of thickness is considerably improved if the low-level sensible heating is included. In studying the initial-value problem for the Eady model, Farrell (1984) indicated that energy extracted from the mean flow during the initial development of a perturbation is able to excite persistent normal modes. It is suggested that this process may be important to cyclogenesis and in providing energy to neutral or near-neutral normal modes. In particular, Farrell had shown that Petterssen’s mechanism can lead to appreciable cyclogenesis. However, in real cyclogenesis events, the initial state is often disturbed by a localized forcing instead of a periodic forcing as assumed by the normal mode method. Observations of East Coast cyclogenesis indicate that low-level diabatic heating is essential in forming the surface low (e.g., Bosart 1981, 1988). Thus, it is important to study the response of a baroclinic flow to localized, low-level sensible heating associated with East Coast cyclogenesis events.

The mathematical problem of prescribed diabatic heating in a mesoscale flow has been shown to be useful in understanding the dynamics of different mesoscale phenomena by several authors (e.g., Smith and Lin 1982; Lin and Smith 1986; Lin 1986; Raymond 1986). Using a linear theory of sheared flow over a diabatic heating region, Lin (1987) found that the low-level upward motion may be located over the warm or cold side of the coastline depending upon the Richardson number associated with the basic flow and the depth of the heating layer. In a separate study, Lin (1989) examined the response of a quasi-geostrophic uniform flow over an isolated warm region. The response is different from that of a smaller-scale flow with rotational effects ignored (e.g., Smith and Lin 1982). Therefore, in order to apply the above work of Lin to the problem of East Coast cyclogenesis, both the rotational and shear effects, i.e., the baroclinic effects, should be included.

A related mathematical problem is the lee cyclogenesis problem which has received considerable attention after the completion of the Alpine Experiment (ALPEX). It is found (e.g., Smith 1984, 1986; hereafter, S84, S86; Speranza et al. 1985; Pierrehumbert 1985) that the baroclinic current is strongly modified by the orography in most cyclogenesis events. That is, the lee cyclogenesis is the result of baroclinic energy conversion. By analogy, it is tempting to hypothesize that the baroclinic energy conversion may also occur for a flow over a region of diabatic heating. However, the response of a stratified flow to thermal forcing is different from that to orographic forcing (Smith and Lin 1982; Lin and Smith 1986). Thus, it is necessary to investigate the response of a baroclinic flow over a region of diabatic heating.

The purpose of this paper is to study the effect of diabatic heating on a baroclinic current with application to the dynamics of East Coast cyclogenesis. In section 2, we describe the theory and the mathematical methods for solving the problem. The theory will be based on quasi-geostrophic equations. Baroclinic wave generation by two-dimensional diabatic heating will be investigated in section 3. Mechanism of the baroclinic wave generation in the vicinity of the diabatic heating region will be discussed. The three-dimensional
response will be studied in section 4. Both circular and elongated heat sources will be considered. The theory will then be applied to the cyclogenesis along the East Coast in section 5. Concluding remarks can be found in section 6.

2. Theory

For an inviscid Boussinesq fluid on an f-plane with constant basic state stratification, the linearized quasi-geostrophic potential vorticity equation and the thermodynamic equation applied at the surface can be written (e.g., S84; Bannon 1986)

\[
\begin{align*}
\left( \frac{\partial}{\partial t} + U \frac{\partial}{\partial x} + V \frac{\partial}{\partial y} \right) (\nabla H^2 p + \frac{f^2}{N^2} p_{xx} ) &= \left( \frac{g \rho_0 f^2}{c_p T_0 N^2} \right) q_x \\
\left( \frac{\partial}{\partial t} + U \frac{\partial}{\partial x} + V \frac{\partial}{\partial y} \right) \theta + u_x \partial_x + v_x \partial_y + w \theta_x \\
&= \left( \frac{\theta_0}{c_p T_0} \right) q \quad \text{at} \quad z = 0,
\end{align*}
\]

where

\( p \) perturbation pressure
\( \theta \) perturbation potential temperature
\( q \) diabatic heating rate per unit mass
\( u_x \) perturbation geostrophic velocity in x direction
\( v_x \) perturbation geostrophic velocity in y direction
\( w \) vertical velocity
\( U \) basic wind velocity in x direction
\( V \) basic wind velocity in y direction
\( \theta \) basic state potential temperature
\( \theta_0 \) constant characteristic potential temperature
\( \rho_0 \) basic state density (constant)
\( T_0 \) basic state temperature
\( N \) Brunt–Väisälä frequency
\( f \) Coriolis parameter
\( g \) gravitational acceleration
\( c_p \) specific heat capacity at constant pressure

and where subscripts in \( u_x \) and \( v_x \) denote partial differentiation.

With the hydrostatic, geostrophic wind and thermal wind equations,

\[
\begin{align*}
\theta &= (\theta_0 / \rho_0) p_x, \\
u_x &= (-1 / f \theta_0) p_y, \\
u_x &= (1 / f \rho_0) p_x, \\
U_x &= (-g / f \theta_0) \partial_y, \\
V_x &= (g / f \theta_0) \partial_x.
\end{align*}
\]

Equation (2) becomes

\[
\begin{align*}
\left( \frac{\partial}{\partial t} + U \frac{\partial}{\partial x} + V \frac{\partial}{\partial y} \right) p_x - (U_x p_x + V_x p_y) \\
+ \rho_0 N^2 w = \left( \frac{g \rho_0}{c_p T_0} \right) q \quad \text{at} \quad z = 0.
\end{align*}
\]

The baroclinic waves associated with the system of Eqs. (1) and (6) are dispersive waves with real frequencies (S84), which can propagate along the surface of the earth in the presence of a horizontal temperature gradient.

Usually the low-level sensible heating/cooling is concentrated in the boundary layer, which has a height of about 1 to 1.5 km. The deformation depth or Rossby height of the flow, i.e. \( H_0 = fL/N \), has a value of about 10 km for a flow with \( f = 10^{-4} \) s\(^{-1}\), \( L \) (horizontal scale) = 1000 km, and \( N = 10^{-2} \) s\(^{-1}\). Compared with the deformation depth, the thickness of the diabatic heating is very small. In this way, we may assume there exists no interior thermal forcing as a first approximation. Notice that the convective latent heating is ignored in this study, which may play an important role in the deepening process of the cyclogenesis. A similar approach has been taken by Bannon (1984) in studying the surface frontogenesis. Thus Eq. (1) reduces to a homogeneous form,

\[
\nabla H^2 p + \left( \frac{f^2}{N^2} \right) p_{xx} = 0.
\]

Making the double Fourier transform in \( x \) (\( \leftrightarrow k \)) and \( y \) (\( \leftrightarrow l \)) of the above equation gives

\[
\hat{p}_{xx} - \left( \frac{N \kappa}{f} \right)^2 \hat{p} = 0,
\]

where \( \kappa = (k^2 + l^2)^{1/2} \) is the horizontal wavenumber. The general solution of Eq. (8) can be written

\[
\hat{p}(k, l, z, t) = A_1(k, l, t)e^{-N |\kappa| z / f} + A_2(k, l, t)e^{N |\kappa| z / f}.
\]

The upper boundary condition requires the solution to vanish at infinity, which implies that \( A_2 = 0 \). The lower boundary condition requires \( w = 0 \) at \( z = 0 \) for a flow over a flat surface such as all cases considered in this paper. Applying the lower boundary condition to Eq. (6) and making Fourier transform in \( x \) and \( y \), we obtain

\[
\hat{p}_{zz} + (ikU_0 + ilV_0) \hat{p}_z - (ikU_x + ilV_x) \hat{p} = \left( \frac{g \rho_0}{c_p T_0} \right) \hat{q},
\]

where \( U_0 \) and \( V_0 \) are the surface wind speeds in \( x \) and \( y \) directions, respectively. The vertical shears, \( U_x \) and \( V_x \), are assumed to be constant throughout this paper. Substituting Eq. (9) into (10) we obtain,

\[
\hat{p}_z + \left[ ik \left( U_0 + \frac{f U_x}{N |\kappa|} \right) + il \left( V_0 + \frac{f V_x}{N |\kappa|} \right) \right] \hat{p} = \left( \frac{-g \rho_0 f}{c_p T_0 N |\kappa|} \right) \hat{q}.
\]

The above equation is similar to Eq. (4.1) of S84 except for the forcing term. For a flow over a steady
state warm/cold region described by the surface potential temperature anomaly, \( T(x, y) \), the heating rate can be calculated according to the following formula (Stern and Malkus 1953):

\[
q(x, y) = c_p \left( U_0 \frac{\partial}{\partial x} + V_0 \frac{\partial}{\partial y} \right) T(x, y). \tag{12}
\]

Making the Fourier transform of (12) and a straightforward manipulation of Eq. (4.1) of S84 and Eq. (11), we obtain a relationship between the orographic forcing and the thermal forcing, namely,

\[
h(x, y) = \left( -\frac{g}{T_0 N^2} \right) T(x, y), \tag{13}
\]

where \( h(x, y) \) is the shape of the mountain. Equation (13) means that the response of a quasi-geostrophic flow over a stationary cold (warm) region is equivalent to that over a mountain (valley) if the forcings are of the same shape. According to the above equation, a cold region with potential temperature anomaly of 5.3 K corresponds to a mountain with height of 2 km if \( T_0 = 260 \) K and \( N = 0.01 \) s\(^{-1}\). This analogy has also been illustrated by Smith (1979, Fig. 15) in which an anticyclonic circulation can be produced by a quasi-geostrophic flow over either a mountain or a cold dome.

The general solution of Eq. (11) can be written as

\[
\hat{p}(k, l, z, t) = Ae^{-Bt} - \left( \frac{-g \rho_0 f}{c_p T_0 N |k|} \right) \frac{\hat{q}(k)}{B} \tag{14}
\]

where \( B \) is defined as

\[
B = i k \left( U_0 + \frac{f U_z}{N |k|} \right) + il \left( V_0 + \frac{f V_z}{N |k|} \right). \tag{15}
\]

The coefficient \( A \) can be determined by the initial condition which is assumed to be

\[
p(t = 0) = 0. \tag{16}
\]

Solving for \( A \) and making use of the Fourier transform of Eq. (12) gives

\[
\hat{p}(k, l, z, t) = \left( \frac{-g \rho_0 f}{c_p T_0 N |k|} \right) \hat{q}(k,l)(1 - e^{-Bt})e^{-N|k|z/f} \frac{1}{B} \tag{17}
\]

The perturbation pressure in the physical domain is then recovered by the inverse Fourier transform

\[
p(x, y, z, t) = \int_{-\infty}^{\infty} \hat{p}(k, l, z, t) e^{i(kx+ly)} dk dl. \tag{18}
\]

Equations (17) and (18) describe the formation of a baroclinic cyclone if there exists a level in which the basic wind reverses direction (as will be discussed later). This is similar to the lee cyclogenesis problem as studied in S84 and S86. We thus propose this as a possible prototype of East Coast cyclogenesis.

### 3. Baroclinic wave generation by two-dimensional diabatic heating

For a two-dimensional quasi-geostrophic flow with diabatic heating, Eqs. (17) and (18) reduce to

\[
p(x, z, t) = \int_{-\infty}^{\infty} \left( \frac{-g \rho_0 f}{c_p T_0 N} \right) \hat{q}(k) \frac{(1 - e^{-Bt})e^{-N|k|z/f}e^{ikx}}{|k|B} \tag{19}
\]

where

\[
B = ik(U_0 + Hu_z); \quad H = f/N |k|. \tag{20}
\]

As discussed in S84, the integral in Eq. (19) will go to zero \((p \rightarrow 0)\) as \(|x| \rightarrow \infty\) due to the rapid oscillation of the \(e^{ikx}\) term if the integrand is well behaved according to the Riemann–Lebesque lemma (Lighthill 1970). This implies that the disturbance will remain locally in the vicinity of the diabatic heat source/sink. The baroclinic waves can only be generated if the denominator of the integrand vanishes for some value of \( k \). This is possible if \( U_0 \) and \( U_z \) have opposite signs, i.e., if there exists a backstreaming basic flow and a wind reversal level \(|k^*| = f/NH^* = -fU_z/NU_0\). An asymptotic solution for large \( x \) and \( t \), similar to that of S86, can be obtained, which describes a train of baroclinic waves extending from the center of the diabatic heating to a moving point \( x = Ud \).

One useful function that may be chosen for a diabatic heating is

\[
q(x) = \frac{Q_0}{1 + (x/a)^2}. \tag{21}
\]

The above equation describes a bell-shaped heat source/sink with a horizontal scale of \( a \). The Fourier transform of \( q(x) \) is

\[
\hat{q}(k) = Q_0 a e^{-|k|a}. \tag{22}
\]

A fast Fourier transform (FFT) algorithm is then employed to solve Eqs. (19) and (22).

Figure 1 shows an example of a baroclinic quasi-geostrophic flow over a diabatic cooling region as described by Eq. (21) with \( Q_0 = -0.24 \) J (kg-s\(^{-1}\)) and \( a = 75 \) km. The basic wind is assumed to be \( U(z) = (-10 + 0.005z) \) m s\(^{-1}\). This gives a wind reversal level of \( H = 2 \) km. The grid interval and the number of grid points in \( x \) direction used in the calculation are 30 and 128 km, respectively. After 6 h (Fig. 1a), there exists a region of perturbation high pressure near the center of diabatic cooling \((x = 0)\). The high is associated hydrostatically with cold air near the cooling center. On the downstream side \((x < 0)\), there exists a wider region of weak low-pressure perturbation. The disturbance decays exponentially with height as indi-
Fig. 1. Two-dimensional baroclinic waves forced by diabatic cooling described by Eq. (21) with $Q_0 = -0.24 J (kg s)^{-1}$ and $a = 75$ km. The basic wind is $U(z) = (-10 + 0.005z) m s^{-1}$. Other parameters are assumed to be $f = 10^{-5} s^{-1}$, $N = 10^{-2} s^{-1}$, $T_0 = 260 K$, and $\rho_0 = 1 kg m^{-3}$. Six levels and four time steps of perturbation pressures are shown: (a) 6 h, (b) 12 h, (c) 18 h, and (d) 24 h. The level of wind reversal is located at 2 km (labeled by $z = H$). The location of an air parcel, originating at $x = 0$ and moving with the group velocity ($c_g = U_0 = -10 m s^{-1}$) is indicated by a dot at each time step. The arrows in (a) illustrate the direction of the basic state wind vectors. The dashed line in (d) is the phase line.

cated by Eq. (19). After 12 h (Fig. 1b), the perturbation high pressure strengthens to a value of about 3.4 mb, while the perturbation low pressure deepens gently to a value of about −1.5 mb. After 18 h (Fig. 1c), the high pressure deepens to about 3.6 mb, while the low pressure increases to about −3.8 mb. After 24 h (Fig. 1d), the high pressure weakens to about 3.2 mb, while the low pressure keeps strengthening to a value of −5.6 mb. The upshear vertical tilt of the trough is evidence of the baroclinic wave generated by the diabatic heating. This allows the heat flux to be transported in the positive y direction (e.g., see Gill 1982). The phase line of the trough becomes more vertical at later stage (not shown). Once the available potential energy (APE) stored in the basic baroclinic current has been transferred to the forced baroclinic waves, the phase line will become vertical. For the present case with $H = 2$ km (wind reversal level), the theory predicts a reason-
able wavelength of 1250 km ($\lambda = 2\pi/k^* = 2\pi NH/f$) of the baroclinic wave with a dipolar structure.

Figure 2 shows the time evolution of the absolute minimum and maximum surface perturbation pressures for the case of Fig. 1. The perturbation high pressure grows rather rapidly in the early stage, reaches its maximum of 3.65 mb at 17 h, and then decays gradually afterwards. The perturbation low pressure develops rather slowly in the first 12 h, then deepens much more rapidly at later stage. The rapid development of the perturbation low pressure after 12 h can be explained by a group velocity argument. The group velocity of the baroclinic wave (S84, S86) is

$$c_g = U(H) - \frac{f}{N|k|^3}(k \cdot U_z)k,$$

where $H = f/N|k|$ and $U(z) = U_0 + U_z z$. The above equation reduces to $c_{gx} = U_0$ for a two-dimensional wave (S84). As indicated in Eq. (12), a moving airstream over the diabatic cooling, $q(x) = Q_0/[1 + (x/a)^2]$, corresponds to that over a cold region extending from $x = 0$ to $-\infty$, $T(x) = (Q_0a/c_pU_0)\tan^{-1}(x/a)$. This is analogous to an airflow over a flat plain from a plateau, according to Eq. (13). Therefore, the fluid is trying to form a high in the vicinity of the cooling center ($x = 0$) and a first trough downstream ($x < 0$). Similar results have been found in S84. For example, consider an air parcel originating at $x = 0$ near the surface (denoted by a dot in Fig. 1). It will take 12 h to advect to 432 km (i.e., $-432$ km in the figure). Downstream at the group velocity $c_g = U_0 = -10$ m s$^{-1}$, which is about the region of the developing low (Fig. 1b). During the 12 to 24 h period, the air parcel reaches the region of the developing low. Thus the low deepens much more rapidly at this stage (Figs. 1c and 2). Like the perturbation high pressure near the cooling center, the perturbation low pressure will reach a minimum and increase its amplitude afterwards since the air parcel originating at $x = 0$ near the surface will pass through the region of the well-developed low. Thus we may conclude that the diabatic heating plays an important role in converting the APE stored in the baroclinic current to the forced baroclinic wave.

Figure 3 shows a case similar to that of Figs. 1 and 2 except for diabatic heating, which corresponds to a cold air advection according to Eq. (12). The dipolar structure of the forced baroclinic waves is also evident, but with the positions of low and high pressure reversed. Similar to the previous case, the low is associated hydrostatically with warm air near the heating center. The time evolution of absolute minimum and maximum surface perturbation pressures for this case is the same as Fig. 2 but with curves a and b reversed.

To show the importance of the baroclinicity and the existence of the wind-reversal level in the above cyclogenesis mechanism, we perform four similar cases except with neither baroclinicity nor a wind-reversal level (Fig. 4). For quasi-geostrophic, baroclinic flow over the diabatic heat source with forward shear (i.e., no wind reversal, Figs. 4a and b), the disturbance is much weaker compared with the corresponding cases with wind reversal (Figs. 1 and 3). This indicates that forward vertical wind shear tends to suppress the development of the low or high pressure. For quasi-geostrophic, barotropic flow over the diabatic heating (cooling) region, a perturbation low (high) of $-5$ mb (+5 mb) is produced after 24 h (Fig. 4c and d). The surface low (high) produced by the diabatic heating (cooling) is located about 400 km downstream of the heating (cooling) center. Notice that a moving airstream over the diabatic heating corresponds to that over a warm region extending from $x = 0$ to $-\infty$ for Fig. 4c. The low pressure at the surface is produced by the less dense air above the warm region in a barotropic flow as required by the hydrostatic equation. Thus the low pressure forms on the warm side or the downstream side of the diabatic heating center ($x = 0$). The results are consistent with the quasi-geostrophic flow over a warm region as studied by Lin (1989). The response is quite different from the low-high couplet produced by diabatic heating in a backsheared baroclinic flow.

4. Three-dimensional response

Two cases of three-dimensional flow over a diabatic heat source are presented in this section. The heat source of case 1 (C-1) is assumed to be circular, while it is elongated in case 2 (C-2). Figures 5 and 6 show the perturbation fields of a baroclinic flow over a bell-shaped heat source with circular contours for C-1. The diabatic heating function is described as

$$q(x, y) = \frac{Q_0}{[x^2/a_x^2 + y^2/a_y^2 + 1]^{3/2}},$$

FIG. 2. Time evolutions of absolute minimum (a) and maximum (b) surface perturbation pressures for the case of Fig. 1.
where $a_x$ and $a_y$ are the horizontal scales of the heat source in $x$ and $y$ directions, respectively. The maximum heating rate ($Q_0$) and the horizontal scale ($a_x = a_y$ in this case) of the thermal forcing are 0.24 J (kg-s)$^{-1}$ and 150 km, respectively. The basic wind is assumed to be unidirectional with $U = (-10 + 0.005z)$ m s$^{-1}$ and $V(z) = 0$ m s$^{-1}$. The basic wind blows from east at the surface and reverses its direction at $z = 2$ km (H). The solution of perturbation pressure is obtained by solving Eq. (17) numerically using a FFT. The perturbation potential temperature ($\bar{\theta}$), relative vorticity ($\bar{\zeta}$), and streamlines are calculated from the following relationship:

$$\bar{\theta} = -\left( \frac{\theta_0}{g \rho_0} \right) \left( \frac{N}{f} \right) \hat{p}, \quad (24)$$

$$\bar{\zeta} = -\left( \frac{\kappa^2}{f \rho_0} \right) \hat{p}, \quad (25)$$

$$\bar{u}_x = -\left( \frac{il}{f \rho_0} \right) \hat{p}, \quad (26)$$

$$\bar{v}_y = \left( \frac{ik}{f \rho_0} \right) \hat{p}, \quad (27)$$
in the Fourier space and then transformed back to the physical space by FFT.

In response to this isolated diabatic heating, a region of low pressure with a minimum of about \(-3.35\) mb forms in the vicinity of the heating after 12 h (Fig. 5a). There exist a broad region of relatively weak high pressure upstream and a compact region of stronger high pressure downstream of the concentrated heating region. Associated with the perturbation low and high pressures are more compact regions of warm and cold air, respectively (Fig. 5b). This is consistent with the hydrostatic balance as required by Eq. (3). The low-high couplet associated with the forced baroclinic wave is located in the vicinity of the forcing region. This is consistent with that implied by Eqs. (19) and (20); the disturbance remains locally in the vicinity of the thermal forcing because the forced baroclinic wave has a zero phase speed (S84). At the surface, the fluid parcel experiences a cyclonic circulation near the center of the low pressure region where there exists a cell of pos-
itive relative vorticity (Fig. 5c). A region of very weak negative vorticity is generated upstream of the heating, while a region of stronger negative vorticity is generated downstream of the heating. An inverted trough forms near the heating center (Fig. 5d), which is often observed to form along the Carolina coast in many cyclogenesis events.

Figure 6 shows the surface fields of perturbation pressure, perturbation potential temperature, relative vorticity, and streamlines and vertical cross sections of perturbation pressure and potential temperature for C-I after 24 h. The forced baroclinic waves are still developing at this stage, as indicated in the time evolution of the low pressure and the maximum perturbation...
potential temperature (Fig. 7). The low reaches a minimum of about $-4.6$ mb, while the high located downstream reaches a maximum of about $2.8$ mb. The high pressure upstream of the heating center remains weak (Fig. 6a). The perturbation potential temperature reaches a maximum of $9.6^\circ C$ and a minimum of $-4^\circ C$ after 24 h (Figs. 6b and 7b). The air parcel undergoes a stronger cyclonic circulation near the heating center and anticyclonic circulation downstream (Fig. 6c) compared with 12 h before. The low remains in the vicinity of the diabatic heating. One interesting feature of the streamline field (Fig. 6d) is that an inverted ridge forms downstream of the low at about $x = -500$ km. It is often observed during East Coast cyclogenesis events (e.g., Richwein 1980; Forbes et al. 1987) that an inverted ridge forms just east of the Appalachians at about the same time the inverted trough forms along the Carolina coast. It has been assumed by many au-
thors that this ridge is the result of cold air damming associated with mountains. The present results suggest than an inverted ridge can form over flat terrain in response to surface heating and vertical wind shear alone. Indeed, the cross section of perturbation potential temperature along $y = 0$ (Fig. 6f) indicates that a pool of cold air is located to the west of the low. The cross section of perturbation pressure along $y = 0$ shows a similar structure as the two-dimensional case (Fig. 3d), but with a more compact low. Both vertical cross sections of pressure and temperature fields (Figs. 6e and 6f) indicate that the disturbance is confined in a shallow layer.

Figure 7a shows the time evolution of the absolute minimum of the perturbation pressure. The low pressure develops more rapidly in the first 10 h because of the direct heating underneath. After 10 h, the low is still developing but at a slower rate. The low pressure reaches a minimum of about $-4.6 \, \text{mb}$ after 24 h. The time evolution of the maximum perturbation potential temperature is depicted in Fig. 7b. The air is heated up more rapidly in the early stage at a rate of about
0.67°C h⁻¹ in the first 10 h. After 10 h, the air is heated up at a slower rate. Both pressure and temperature fields indicate that the low is still developing at 24 h.

Figure 8 shows a case (C-2) similar to C-1 except the heat source is elongated four times wider in the y direction, i.e. $a_x = 75$ km and $a_y = 300$ km in Eq. (23). The surface fields of perturbation pressure, perturbation potential temperature, and relative vorticity at 12 h appear to have a dipolar pattern in the basic flow direction but elongated in y direction (Figs. 8a–c). The result indicates that the horizontal pattern of the disturbance is directly related to the geometry of the forcing. Similar results have also been found in a nonrotational, uniform flow over an elevated heat source (Lin 1986). The streamline field (Fig. 8d) depicts that an inverted trough is developed near the elongated heat source. The forced baroclinic wave strengthens after 24 h. The patterns of the low, warm

---

**Fig. 8.** As in Fig. 5 except with $a_x = 75$ km and $a_y = 300$ km (C-2).
region, and positive vorticity region near the heating center become more circular after 24 h, while the patterns of the high, cold region, and negative relative vorticity region still keep the elongated shape as 12 h before (Figs. 9a–c). A weak positive vorticity forms downstream of the region of negative vorticity at this stage (Fig. 9c). The streamline field (Fig. 9d) shows that a well-developed inverted ridge forms downstream of the inverted trough, compared with C-1 (Fig. 6d).

5. Application to East Coast cyclogenesis

As mentioned in Introduction, a type A cyclone forms as a wave on the cold front associated with a cold synoptic-scale anticyclone located to the north of the Carolina coast. Under such a synoptic setting, the basic wind blows roughly from northeast near the surface and becomes a southwesterly in the upper layer. The 1979 Presidents' Day snowstorm (Bosart 1981;
Fig. 10. Surface analysis for 0000 UTC 25 January 1986 (GALE IOP 2). The horizontal winds and isotherms (dotted), and the sensible heat fluxes (in W m$^{-2}$) are shown in (a) and (b), respectively. Full bar represents 5 m s$^{-1}$, and half bar represents 2.5 m s$^{-1}$. The square dots represent the PAM sites. The dashed lines in (b) represent the estimated values of sensible heat flux where only sparse data are available (provided by A. J. Riordan).
Uccellini et al. (1984) and the cyclogenesis during the second GALE IOP 2 (GALE 1986) are good examples. Three simulations have been performed in this section. In case 3 (C-3), the model is applied to the East Coast cyclogenesis of GALE IOP 2, in which the surface wind blows from 60° and the heat source is elongated and oriented in the northeast–southwest direction. Case 4 (C-4) is similar to C-3 but with an easterly surface wind.

**FIG. 11.** A developing low produced by a baroclinic current over a diabatic heating region after 12 h (C-3). The diabatic heating function is prescribed by Eq. (28). The parameters used are $Q_0 = 0.24 \, J \, (kg \cdot s)^{-1}$, $a_x = 75 \, km$, and $a_y = 300 \, km$. The basic wind is $U(z) = (-8.7 + 0.00415z) \, m \cdot s^{-1}$ and $V(z) = (-5 + 0.00347z) \, m \cdot s^{-1}$. The basic wind blows from 60° with a speed of 10 $m \cdot s^{-1}$ at surface and 225° with a speed of 20 $m \cdot s^{-1}$ at 500 mb. The contour of heating rate of 0.04 $J \, (kg \cdot s)^{-1}$ (bold) and the hodograph are shown in (a). $U_0$, $U_z$, and $U_T$ denote the surface, 500 mb (5.5 km) and thermal winds, respectively. Four surface fields illustrated are: (a) perturbation pressure, (b) perturbation potential temperature, (c) relative vorticity, and (d) streamlines.
Fig. 12. As in Fig. 11 except for 24 h (C-3). Vertical cross sections of the perturbation pressure and perturbation potential temperature along $y = 0$ are shown in (e) and (f).

wind. To show the importance of the baroclinicity in a three-dimensional flow over a diabatic heat source, a barotropic case (C-5) is performed.

To apply the above theory to East Coast cyclogenesis, the prescribed heating is specified with observed values of sensible heat flux. Figure 10 shows the surface analysis of horizontal wind, temperatures, and sensible heat flux near the Carolina coast for 0000 UTC 25 January 1986, just before the cyclogenesis event of GALE IOP 2 (GALE 1986). A region of positive sensible heat flux with a maximum of about $240 \text{ W m}^{-2}$ is located near the western boundary of the Gulf Stream. As can be seen from the surface temperature distribution and the surface wind field, the diabatic heating is associated
with the cold air advection. For simplicity, we rotate
the diabatic heating function, Eq. (23), clockwise by
45 degrees to represent this surface sensible heating,

\[ q(x, y) = \frac{Q_0}{[(x - y)^2/2a_x^2 + (x + y)^2/2a_y^2 + 1]^{3/2}}. \]  

(28)

The same horizontal scales, \(a_x = 75\) km and \(a_y = 300\)
km, as C-2 have been used. The maximum heating
rate, \(0.24\) J (kg-s\(^{-1}\)), is roughly estimated by dividing
the maximum sensible heat flux (240 W/m\(^2\)) by \(\rho_0 D_\theta\),
where \(\rho_0\) is the constant basic state density (1 kg m\(^{-3}\))
and \(D_\theta\) is the depth of the heating (1 km). The basic
wind is assumed to be \(U(z) = (-8.7 + 0.00415z)\) m
s\(^{-1}\) and \(V'(z) = (5.0 + 0.00347z)\) m s\(^{-1}\).

At the surface, the basic wind blows from 60° with a speed of 10 m s\(^{-1}\), which is estimated from Fig. 10a. At 500 mb (5.5 km), it blows from 225° (southwest) with a speed of 20 m s\(^{-1}\), which is roughly estimated from the 500 mb chart (GALE 1986). The thermal wind is almost parallel to the coastline. The hodograph of the basic state wind vectors at surface and 500 mb is sketched in the upper left corner of Fig. 11a.

Figure 11 shows the surface fields of perturbation pressure, perturbation potential temperature, relative vorticity, and streamlines after 12 h. A surface low with an amplitude of −3.6 mb develops initially near the center of diabatic heating. The genesis region of the
low is near the western boundary of the Gulf Stream, which is consistent with observations of GALE IOP 2 (GALE 1986) and most of type A cyclogenesis (Colucci 1976). The low is elongated in the northeast–southwest direction as is the heat source. In contrast to C-2, there is no coupled high pressure produced at this time. Associated with the low pressure is a more compact region of warm air and positive relative vorticity (Figs. 11b and 11c). The maximum perturbation potential temperature and the positive relative vorticity reaches 8.5°C and 3.9 × 10⁻⁴ s⁻¹. A region of weak cold air and negative relative vorticity has already been produced to the northwest of the warm air and positive relative vorticity. A cyclonic circulation associated with the positive vorticity can also be found near the low pressure center (Fig. 11d). The disturbance is stronger compared with C-2 (Fig. 8). This is due to more heat received by the surface wind (60°) near the surface.

In the next 12 h period (12–24 h, Fig. 12), the low grows at a slower rate compared with the first 12 h period (Fig. 13). The low reaches a minimum of −5.3 mb, cuts off from the flow and still keeps growing at 24 h. For the entire period, the cutoff low remains in the vicinity of the diabatic heat source. During this period, two regions of high pressure are developed, with the stronger (weaker) one located to the west (south) of the low (Fig. 12a). Both high pressures are associated with the cold air and negative relative vorticity (Figs. 12b and 12c). An air parcel near the surface upstream of the heating region undergoes a strong cyclonic circulation near the low (Fig. 12d). The inverted ridge to the west of the low is weaker compared with C-1 and C-2 (Figs. 6d and 9d). However, a pool of cold air can still be found to the west of the low from the vertical cross section of the perturbation potential temperature along y = 0 (Fig. 12f). A strong anticyclonic flow forms to the south of the cyclone, which is associated with the negative relative vorticity located at the southeast of the low (Fig. 12d). One interesting feature of the streamline field is that a confluent zone is produced to the northeast of the low, while a diffusent zone is produced to the southwest. The confluent–diffusent couplet may be related to that associated with coastal front which often precedes the cyclogenesis (Riordan 1988). Similar to C-1 and C-2, the disturbance is confined in a shallow layer as indicated by the cross sections of the perturbation pressure and perturbation potential temperature along y = 0 (Figs. 12e and 12f). Some weak disturbances can be found at the northwest, southwest, and southeast corners, which are generated by the periodic nature of the FFT algorithm. Beyond 24 h, the strength of the cyclone and the speed of the northeasterward movement, as often observed, may be strongly influenced by the moisture effect and upper-level forcing such as jet streaks.

Figure 14 shows the flow response at 24 h of a case (C-4) similar to C-3 but with \( U(z) = (-10 + 0.005z) \) m s⁻¹ and \( V(z) = 0.004z \) m s⁻¹. The basic state wind at the surface is from east. The basic characteristics of the flow response, such as the formation of cyclone and the trough–ridge couplet are similar to that of C-3. The inverted trough and ridge are oriented in the northeast–southwest direction. The low reaches a minimum of −2.8 mb, which is weaker than that of C-3. This is due to lesser heat received by the easterly wind at surface compared with the northeasterly wind of C-3. The high reaches a maximum of 1.4 mb, which is slightly higher than that in C-3. The low moves to a distance of about 400 km north of the genesis region. This may be due to a stronger northward component.
of the basic state wind compared with C-3 according to the group velocity argument (S86). Regions of warm air and positive vorticity are asymmetric with respect to the southwest–northeast direction (60°) with the maxima located near the heating center. The low starts to decay after 17 h (Fig. 14e) because it has moved away from the concentrated heating area. The perturbation potential temperature at the surface (Fig. 14f) reaches its maximum of 5.1°C at 13 h and then decreases. Notice that the confluent–diffluent zones are less pronounced and the inverted ridge is more pronounced compared with C-3.

Figure 15 shows a barotropic flow over a diabatic heat source. The basic wind is assumed to be from southwest of 5 m s⁻¹ (\( U = V = 3.536 \) m s⁻¹). The response is quite different from the baroclinic cases (C-

**Fig. 15.** As in Fig. 12 except for a barotropic case (C-5). The basic wind blows uniformly from the southwest with a speed of 5 m s⁻¹ (\( U = V = 3.536 \) m s⁻¹).
3 and C-4). The disturbance along the southwest-northeast direction is similar to the two-dimensional barotropic case (Fig. 4c). A low forms near the heating center and moves northeastward. Similar to the nonrotating flow over a prescribed heating, the movement of the disturbance is steered by the basic wind (Lin and Smith 1986). The cutoff low reaches a minimum of about −6.5 mb at 24 h. As discussed in section 3, the low pressure is produced by the less dense air above the warm region as required by the hydrostatic equation. Two regions of weak high are located to the upstream and downstream of the heating center. This is different from the forced baroclinic cases (C-3 and C-4) in which the two highs are oriented in the northwest-southeast direction and the low extends more on the downstream side. A compact region of warm (cold) air and positive (negative) relative vorticity is associated with the low (high). A cyclone circulation associated with the positive vorticity is also evident near the low pressure center (Fig. 15d). However, this type of flow is rarely observed in East Coast cyclogenesis events. Along with baroclinic cases, we may conclude that East Coast cyclones are produced hydrostatically by the less dense air above the heating region with the modification of the baroclinic effects.

6. Concluding remarks

A quasi-geostrophic theory of cyclogenesis forced by a low-level diabatic heating is proposed. The linear governing equation was solved analytically in the Fourier space and transformed back to the physical space numerically using a fast Fourier transform (FFT) algorithm. Both two- and three-dimensional cases were investigated. The heating layer was assumed to be shallow in this study.

The cyclone is considered to be generated by a diabatic heating in a backsheared baroclinic flow. Existence of wind reversal in one direction of the basic flow is an essential criterion to obtain a forced baroclinic wave in the vicinity of the heating region. There exists an analogy for a quasi-geostrophic flow over a mountain and a region of steady state diabatic heating. The relationship is described by $h(x, y) = (-g/T_0N^2)T(x, y)$, where $h(x, y)$ is the mountain shape and $T(x, y)$ is the temperature anomaly.

The response of a backsheared baroclinic flow over a region of two-dimensional diabatic heating (cooling) is a coupled low–high (high–low) pressure pair located in the vicinity and on the downstream side of the heating (cooling), respectively. The pressure perturbation near the heating (cooling) region develops earlier than that on the downstream side. Physically, the growth of the pressure perturbation can be explained by a group velocity argument. The disturbance remains locally in the vicinity of the forcing due to zero phase speed of the forced baroclinic wave. The upshear phase tilt indicates that the disturbance is a manifestation of a forced baroclinic wave.

The response of an east–west backsheared baroclinic flow over an isolated region of diabatic heating with circular contours is a growing cyclone located near the center of the heat source. A coupled high pressure forms downstream of the diabatic heating. The disturbance is confined in a very shallow layer. The low resembles the geometry of the heat source in the early stage, which then strengthens and develops to be more circular at later stage. The high located downstream of the heat source tends to develop at later stage. An inverted trough associated with an elongated heat source is more pronounced. One interesting finding of this study is that an inverted ridge forms downstream of the low or the inverted trough.

When applied to East Coast cyclogenesis, a cyclone develops near the center of the region of maximum diabatic heating, i.e., near the western boundary of the Gulf Stream. The cutoff low remains in the vicinity of the diabatic heat source. Two regions of weaker high pressure form to the southeast and northwest corners of the low. The inverted ridge becomes weak, however, a pool of cold air can still be found to the west of the low. A strong anticyclonic flow forms to the south of the low. One interesting feature of the flow pattern is that a confluent zone forms to the northeast of the low, while a diffluent zone forms to the southwest of the low. The confluent–diffluent couplet may be related to that associated with coastal front, which often precedes the cyclogenesis. The genesis region and the flow pattern of the cyclone predicted by the theory are consistent with observations. With an easterly wind at the surface, the inverted trough–ridge couplet is more pronounced than with a northeasterly wind. The low starts to decay as it moves out of the concentrated heating region. The East Coast cyclones are produced hydrostatically by the less dense air above the heating region with the modification of the baroclinic effects.

In this study, we have idealized the heating layer to be shallow and prescribed the diabatic heating, which may be improved by the addition of an Ekman friction layer in the theory. Nonlinearity and ageostrophic advection also have been ignored in this work, which will be studied by using a semigeostrophic approach in Part II. However, the present theory provides a valuable physical insight into why an initial cyclone may develop in the planetary boundary layer without major jet dynamics, latent heating or orographic forcing. To include all those factors, a fully numerical modeling might be necessary.

Acknowledgments. The author is grateful to Dr. R. B. Smith for his valuable comments on this work and Dr. A. J. Riordan for reviewing the manuscript and providing Fig. 10. The anonymous reviewers’ comments are highly appreciated. The author wishes to thank Mr. C.-Y. Huang for providing plotting pro-
grams. This work was supported by the FRPD Fund of the North Carolina State University under Grant 01044 and the National Science Foundation under Grant ATM-88-07064.

REFERENCES


